



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/736,349	12/14/2000	Elizabeth Adleberg Brodsky	AUS9-2000-0510-US1	8921

7590 07/19/2004

Casimer K. Salys  
International Business Machines Corporation  
Intellectual Property Law Dept., Internal Zip 4054  
11400 Burnet Road  
Austin, TX 78758

EXAMINER

KLINGER, SCOTT M

ART UNIT	PAPER NUMBER
----------	--------------

2153

DATE MAILED: 07/19/2004

2

Please find below and/or attached an Office communication concerning this application or proceeding.

**Office Action Summary**

Application No.

09/736,349

Applicant(s)

BRODSKY ET AL.

Examiner

Scott M. Klinger

Art Unit

2153

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 14 December 2000.
- 2a) ☐ This action is **FINAL**.                      2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-24 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-24 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- |                                                                                                                        |                                                                                         |
|------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)                                            | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                                   | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)             |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____                                                |

Art Unit: 2153

### **DETAILED ACTION**

Claims 1-24 are pending.

Claim 17 contains a typo: "An apparatus for for crawling"

Claim 23 contains a typo: "a processor connected a network"

### ***Priority***

No claim for priority has been made. The effective filing date for subject matter in the application is 14 December 2000.

### ***Claim Rejections - 35 USC § 112***

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claims 2, 3, 5, 7, 13, 15, 21, and 23 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Regarding claims 2 and 3, claim 2 depends on itself rendering the claim indefinite. Claim 3 depends on claim 2.

Regarding claims 5, 7, 13, 15, 21 and 23, the phrase "at least some of the web pages being dynamically generated" renders the claim indefinite because the minimum number of web pages being dynamically generated is unclear. See MPEP § 2173.05(d).

***Claim Rejections - 35 USC § 102***

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

Claims 1, 4, 9, 10, 12, 17, 18, and 20, are rejected under 35 U.S.C. 102(e) as being anticipated by Najork et al. (U.S. Patent Number 6,301,614, hereinafter “Najork”). Najork discloses a system and method for efficient representation of data set addresses in a web crawler. Najork shows,

In referring to claim 1,

- Querying a web site server by a crawler program, wherein at least one page of the web site has a reference for executing by a browser to produce an address for a next page; parsing such a reference from one of the web pages by the crawler program and sending the reference to an applet running in the browser:  
*“The thread then downloads the document corresponding to the URL, and processes the document (162). That processing may include indexing the words in the document so as to make the document accessible via a search engine. However, the only processing of the document that is relevant to the present discussion is that the main procedure identifies URL's in the downloaded document that are candidates for downloading and processing (step 162). Typically, these URL's are found in hypertext links in the document being processed.”* (Najork, col. 4, line 62 – col. 5, line 4)
- Determining the address for the next page by the browser responsive to the reference and sending the address to the crawler:

*“The web crawler thread determines the URL of the next document to be downloaded (step 160), typically by retrieving it from a queue data structure (not shown).”* (Najork, col. 4, lines 59-62)

In referring to claim 4,

- The crawler is programmable to perform particular action sequences for generating the queries to the web server:

*Najork, col. 4, lines 59-62* (see full quote above)

In referring to claim 9,

- First instructions for querying a web site server by a crawler program, wherein at least one page of the web site has a reference for executing by a browser to produce an address for a next page; second instructions for parsing such a reference from one of the web pages by the crawler program and sending the reference to an applet running in the browser:

*Najork, col. 4, line 62 – col. 5, line 4* (see full quote above)

- Third instructions for determining the address for the next page by the browser responsive to the reference and sending the address to the crawler:

*Najork, col. 4, lines 59-62* (see full quote above)

In referring to claim 10,

- The browser being configured to use a certain proxy, and refer to a resolver file for hostname-to-IP-address-resolution, and wherein the web site server has an IP address, the proxy for the browser has a certain IP address, and the resolver file indicates the certain IP address as the IP address for the web site server:

Najork, Fig. 1 shows a domain name system 114 that provides hostname-to-IP-address-resolution

Art Unit: 2153

In referring to claim 12,

- The first instructions comprise instructions for causing the crawler to perform particular action sequences for generating the queries to the web server.

*Najork, col. 4, lines 59-62 (see full quote above)*

In referring to claim 17,

- A processor connected a network:

Najork, Fig. 1 shows a processor 106 connected to a network 110

- A storage device connected to the processor and the network; the storage device is for storing a program for controlling the processor:

Najork, Fig. 1 shows a storage device 118 storing web crawler program 140

- Querying a web site server by the crawler, wherein at least one page of the web site has a reference for executing by the browser to produce an address for a next page; parsing such a reference from one of the web pages and sending the reference to an applet running in the browser:

*Najork, col. 4, line 62 – col. 5, line 4 (see full quote above)*

- Determining the address for the next page by the browser responsive to the reference and sending the address to the crawler:

*Najork, col. 4, lines 59-62 (see full quote above)*

In referring to claim 18,

- The browser being configured to use a certain proxy, and refer to a resolver file for hostname-to-IP-address-resolution, and wherein the web site server has an IP address, the proxy for the browser has a certain IP address, and the resolver file indicates the certain IP address as the IP address for the web site server:

Najork, Fig. 1 shows a domain name system 114 that provides hostname-to-IP-address-resolution

Art Unit: 2153

In referring to claim 20,

- The processor is operative with the program for causing the crawler to perform particular action sequences for generating the queries to the web server:

*Najork, col. 4, lines 59-62 (see full quote above)*

### ***Claim Rejections - 35 USC § 103***

Claims 5-8, 13-16, 21-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Najork in view of Challenger et al. (U.S. Patent Number 6,026,413, hereinafter "Challenger").

In referring to claims 5, 13, and 21, although Najork shows substantial features of the claimed invention, including the method and apparatus of claims 1 and 17 (see 102 rejection above), Najork does not show caching dynamically generated web pages. Nonetheless this feature is well known in the art and would have been an obvious modification to the system disclosed by Najork as evidenced by Challenger.

In analogous art, Challenger discloses determining how changes to underlying data affect cached objects. Challenger shows processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server: Challenger, Fig. 1C shows the caching of dynamically generated web pages and their dependencies.

Given these teachings, a person of ordinary skill in the art would have readily recognized the desirability and advantages of modifying the system of Najork so as to cache dynamically generated web pages, such as taught by Challenger, in order to increase the speed in which previously viewed web pages are accessed.

In referring to claims 6, 14, and 22, Najork in view of Challenger shows,

- The system of claims 5, 13, and 21 (see 103 rejection above)

Art Unit: 2153

- At least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages:

Challenger, Fig. 1C shows the caching of dynamically generated web pages and their dependencies. Said dependencies used to replace the original references to web pages.

---

In referring to claim 7, Najork shows substantial features of the claimed invention, including querying a web site server by a crawler program responsive to references from one web page to another in the web site, wherein the queries are for causing the server to generate web pages, at least one of the web pages being dynamically generated: *Najork, col. 4, line 62 – col. 5, line 4* (see full quote above)

However, Najork does not show caching dynamically generated web pages. Nonetheless this feature is well known in the art and would have been an obvious modification to the system disclosed by Najork as evidenced by Challenger.

In analogous art, Challenger discloses determining how changes to underlying data affect cached objects. Challenger shows processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server: Challenger, Fig. 1C shows the caching of dynamically generated web pages and their dependencies.

Given these teachings, a person of ordinary skill in the art would have readily recognized the desirability and advantages of modifying the system of Najork so as to cache dynamically generated web pages, such as taught by Challenger, in order to increase the speed in which previously viewed web pages are accessed.



Art Unit: 2153

In referring to claim 8, Najork in view of Challenger shows,

- The system of claim 7 (see 103 rejection above)
- At least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages:

Challenger, Fig. 1C shows the caching of dynamically generated web pages and their dependencies. Said dependencies used to replace the original references to web pages.

---

In referring to claim 15, Najork shows substantial features of the claimed invention, including first instructions for querying a web site server by a crawler program responsive to references from one web page to another in the web site, wherein the queries are for causing the server to generate web pages, at least one of the web pages being dynamically generated: *Najork, col. 4, line 62 – col. 5, line 4* (see full quote above)

However, Najork does not show caching dynamically generated web pages. Nonetheless this feature is well known in the art and would have been an obvious modification to the system disclosed by Najork as evidenced by Challenger.

In analogous art, Challenger discloses determining how changes to underlying data affect cached objects. Challenger shows instructions for processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server: Challenger, Fig. 1C shows the caching of dynamically generated web pages and their dependencies.

Given these teachings, a person of ordinary skill in the art would have readily recognized the desirability and advantages of modifying the system of Najork so as to cache dynamically generated web pages, such as taught by Challenger, in order to increase the speed in which previously viewed web pages are accessed.

Art Unit: 2153

In referring to claim 16, Najork in view of Challenger shows,

- The system of claim 15 (see 103 rejection above)
- At least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and instructions for removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages:

Challenger, Fig. 1C shows the caching of dynamically generated web pages and their dependencies. Said dependencies used to replace the original references to web pages.

---

In referring to claim 23, Najork shows substantial features of the claimed invention, including:

- A processor connected to a network:  
Najork, Fig. 1 shows a processor connected to a network
- A storage device connected to the processor and the network, wherein the storage device is for storing a program for controlling the processor, and wherein the processor is operative with the program to execute a crawler program:  
Najork, Fig. 1 shows a storage device 118 storing web crawler program 140
- A browser program for querying a web site server by the crawler responsive to references from one web page to another in the web site, wherein the queries are for causing the server to generate web pages, at least some of the web pages being dynamically generated; and

*Najork, col. 4, line 62 – col. 5, line 4 (see full quote above)*

However, Najork does not show caching dynamically generated web pages. Nonetheless this feature is well known in the art and would have been an obvious modification to the system disclosed by Najork as evidenced by Challenger.

In analogous art, Challenger discloses determining how changes to underlying data affect cached objects. Challenger shows processing the server generated web pages to generate

Art Unit: 2153

corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server: Challenger, Fig. 1C shows the caching of dynamically generated web pages and their dependencies.

Given these teachings, a person of ordinary skill in the art would have readily recognized the desirability and advantages of modifying the system of Najork so as to cache dynamically generated web pages, such as taught by Challenger, in order to increase the speed in which previously viewed web pages are accessed.

In referring to claim 24, Najork in view of Challenger shows,

- The system of claim 23 (see 103 rejection above)
- At least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages.

Challenger, Fig. 1C shows the caching of dynamically generated web pages and their dependencies. Said dependencies used to replace the original references to web pages.

---

Claims 11 and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Najork in view of Yoshida et al. (U.S. Patent Number 6,748,418, hereinafter "Yoshida"). Although Najork shows substantial features of the claimed invention, including the system of claims 11 and 19 (see 102 rejection above), Najork does not show adding an onload attribute to one of the web pages by the proxy. Nonetheless this feature is well known in the art and would have been an obvious modification to the system disclosed by Najork as evidenced by Yoshida.

In analogous art, Yoshida discloses a technique for permitting collaboration between web browsers and adding content to HTTP messages bound for web browsers. Yoshida shows adding an onload attribute to one of the web pages by the proxy:

Art Unit: 2153

*“The HTTP message editor 123 specifies the script or help HTML to be displayed by referring to the help DB 151 and the script DB 153 based on the HTTP message delivered by 15 the HTTP message checker 125 and the rank and inserts the following program written in JavaScript into the HTTP message.*

```
function openScript(url) {  
    window.open (url, "help_window");  
}  
<body onLoad="openScript (\\"High_Level_Script\\")>  
</body>” (Yoshida, col. 10, lines 52-64)
```

Given these teachings, a person of ordinary skill in the art would have readily recognized the desirability and advantages of modifying the system of Najork so as to add an onload attribute to one of the web pages a proxy, such as taught by Yoshida, in order to allow the web crawler to know when the page is fully loaded.

Art Unit: 2153

*Conclusion*


Any inquiry concerning this communication or earlier communications from the examiner should be directed to Scott M. Klinger whose telephone number is (703) 305-8285. The examiner can normally be reached on M-F 7:00am - 3:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Glenn Burgess can be reached on (703) 305-4792. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Scott M. Klinger  
Examiner  
Art Unit 2153

smk

  
GLENN B. BURGESS  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2100